Korean Journal of Life Science
Vol. 11. No. 5. 400~404, 2001

Identification and Phylogeny of Long Terminal Repeat Elements of Human Endogenous Retrovirus HERV-S

Joo-Young Choi, Joo-Mi Yi, Seung-Heui Jeon, Kyung-Mi Shin, Ji-Won Lee, Won-Ho Lee and Heui-Soo Kim*

Division of Biological Sciences, College of Natural Sciences, Pusan National University, Pusan 609-735, Korea

Abstract

A new human endogenous retroviral family (HERV-S) has recently been identified from human X chromosome. It is 6.7 kb in length and has a typical retroviral structure with LTR-*gag-pol-env*-LTR. Using the PCR and sequencing approach, we investigated LTR elements of the HERV-S family from a human genomic DNA. Four LTR elements (HSL-1, HSL-5, HSL-10, HSL-11) were identified and have a high degree of sequence similarity (96-99%) with that of the HERV-S. Phylogenetic analysis from the HERV-S family indicated that the LTR elements were mainly divided into 2 groups through evolutionary divergence in the primate evolution. Further investigation of the HERV-S LTR elements in primates may cast light on the integration timing into the primate genome and understanding of human evolution.

Key words – HERV-S, Human evolution, LTR elements, Phylogeny

Introduction

Human endogenous retroviruses (HERVs) have been subjected to many amplification and transposition events resulting in a widespread distribution of complete or partial retroviral sequences throughout the human genome. It has been suggested that HERVs have played a role in influencing the functional organization of the human genome[1]. Approximately 1% of the human genome is represented by such sequences. HERVs present full-length or incomplete sequences with multiple stop codons, insertions, deletions, and frameshifts. Such retroviral elements have been described in primates and humans[2,3,10]. Retroviral long terminal repeat (LTR) sequences have the capacity to exert a regulatory influence as promoters and enhancers of cellular genes[8]. Most HERV families encompass a relatively low copy number per haploid genome[7], compared with others that are either high copy number or single-copy retroviral elements[9]. These different copy numbers could either represent multiple integration events or provirus amplified after the integration by retrotransposition.

Recently novel human endogenous retrovirus HERV-S was found using BLAST search[11]. The prototype of the HERV-S was located on cosmid AC004385 that was derived from the human X chromosome. It is 6.7 kb in length and has a typical retroviral structure with LTR-*gag-pol-env*-LTR. We were interested in exploring to identify the LTR elements of HERV-S family in human genomic DNAs.

Materials and Methods

Using the polymerase chain reaction (PCR) approach, we isolated four new retroviral sequences belonging to

*To whom all correspondence should be addressed
Tel : +82-51-581-2962, Fax : +82-51-510-2259
E-mail : khs307@hyowon.cc.pusan.ac.kr

the HERV-S family from a human genomic DNA. New 340-bp LTR elements of HERV-S family were amplified by the primer pair HS47 (5'-TCTGTAAGGAATGT-AGCTGTG-3', bases 54864-54884) and HY77 (5'-CATCCT-GACGACTACACCTA-3', bases 55183-55202) from the HERV-S (GenBank, accession no. AC004385). The PCR conditions followed were those of Kim et al.[5] with an annealing temperature of 54°C. PCR products were separated on a 2% agarose gel, purified with the QIAEX II gel extraction kit (Qiagen) and cloned into the T-khs307 vector[4]. The cloned DNA was isolated by the alkali lysis method using the High Pure plasmid isolation kit (Boehringer Mannheim). Individual plasmid DNAs was screened for inserts by PCR. Positive samples were subjected to sequence analyses on both strands with T7 and M13 reverse primers using an automated DNA sequencer (Model 373A) and the DyeDeoxy terminator kit (Applied Biosystem). Nucleotide sequence analyses were performed using GAP, PILEUP, and PRETTY from the GCG package (University of Wisconsin). Neighbor-joining phylogenetic analysis was performed with the MEGA program [6]. Statistical significance evaluation of the branching pattern was performed with 100 replications.

Results and Discussion

Using the primer pairs derived from the HERV-S LTR (AC004385), we examined to amplify their family in human genome. The PCR primers were designed to hybridize to 5' LTR and not to hybridize to 3 LTR. The observed percentage of divergence of nucleotide sequences between 5' LTR and 3' LTR was 14%. In this study, we examined 5' LTR that contains serine tRNA binding site. To identify HERV-S LTR family, the PCR fragments from the human genomic DNA were cloned and sequenced. Four sequences belonging to the HERV-S were newly identified (Fig. 1) and they showed 96.3~99.0% similarity with that of HERV-S LTR (Table 1). Further, we searched for the LTR elements of the HERV-S family in GenBank database. AC002523 (chromosome Xq28), AL021997 (chromosome 6p22.1~22.3), AL122007 (chromosome 1p13.1~13.3), and AC005678 (chromosome 7) were found by BLAST search. In order to understand the evolutionary relationship within the HERV-S family on human chromosomes, a phylogenetic tree was constructed with the neighbor-joining method using the nucleotide sequences of the LTR elements. As shown in Fig. 2, the HERV-S LTR family was mainly divided into 2 groups through evolutionary divergence in primate genome. Group I contained the all clones (HSL-1, 5, 10, 11) identified newly in this study and also included HERV-S LTR derived from chromosome X. Group II contained AL122007, AL021997, AC005678, AC002523 derived from chromosomes 1, 6, 7, X, respectively. The results

Table1. Percentage similarity of nucleotide sequence of HERV-S LTRs

	1	2	3	4	5	6	7	8	9
1. HERV-S LTR	-								
2. HSL-1	96.3	-							
3. HSL-5	99.0	95.3	-						
4. HSL-10	98.7	94.9	97.6	-					
5. HSL-11	97.0	93.9	95.9	99.0	-				
6. AL122007	86.5	82.4	85.5	85.5	83.7	-			
7. AL021997	87.2	84.1	86.2	85.8	86.1	86.0	-		
8. AC005678	86.9	83.7	85.9	85.5	85.8	85.7	99.7	-	
9. AC002523	84.7	81.5	83.7	84.4	83.8	83.7	88.4	88.0	-

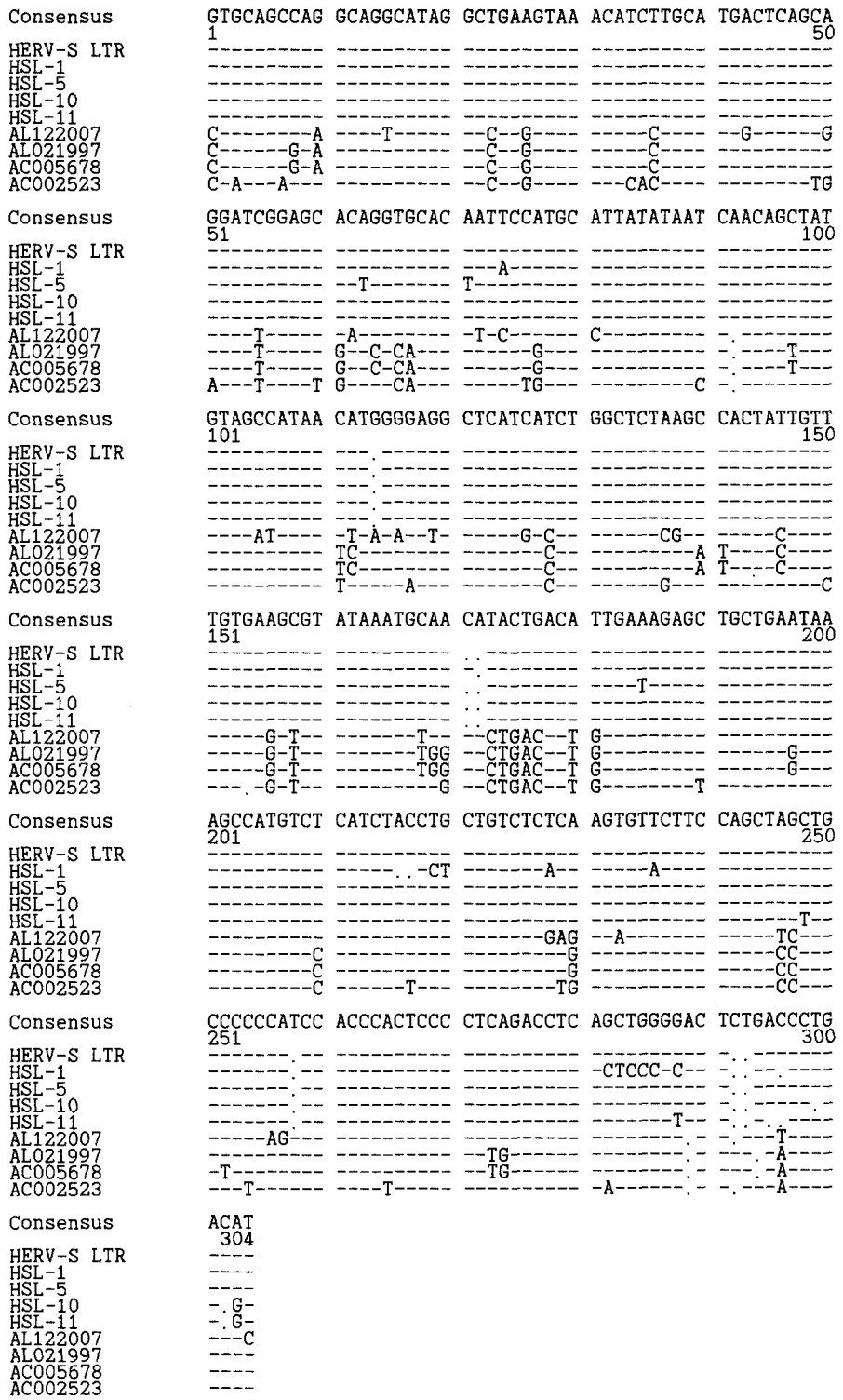

Fig. 1. Sequence alignments of the HERV-S LTR elements.

Consensus sequences are shown in the top row. Dashes indicate no change to the consensus sequences and dots indicate gaps. The nucleotide sequence data reported in this paper will appear in the DDBJ/EMBL/ GenBank nucleotide sequence databases under accession numbers AB051564 (HSL-1), AB051567 (HSL-5), AB051570 (HSL-10), AB051571 (HSL-11).

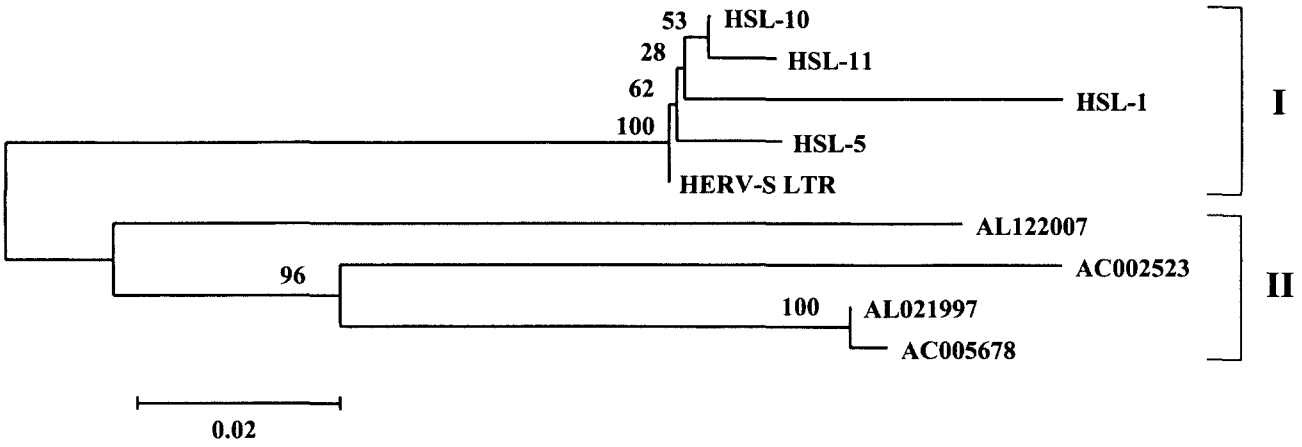

Fig. 2. Phylogenetic tree obtained by neighbor-joining method for the LTR elements of the HERV-S family on human chromosomes. Branch lengths are proportional to the distances between the taxa.
The values at the branch points indicate the percentage support for a particular node after 100 bootstrap replicates were performed.

allowed us to speculate that several copies of HERV-S LTR elements could exist on human genome with at least two different type of HERV-S. Identification of the genes that can be regulated by HERV-S LTR elements on human chromosome will be of great help for further study in understanding of LTR function in relation to human disease and genomic instability in human evolution.

References

1. Baltimore, D. 1985. Retroviruses and retrotransposons: the role of reverse transcription in shaping the eukaryotic genome. *Cell* **40**, 481-482.
2. Kabt, P., M. Tristem, R. Opavsk and J. Pastorek. 1996. Human endogenous retrovirus HC2 is a new member of the S71 retroviral subgroup with a full-length *pol* gene. *Virology* **226**, 83-94.
3. Kim, H. S. and T. J. Crow. 1999. Isolation of novel endogenous retrovirus HC2-like elements in human chromosome. *AIDS Res. Hum. Retroviruses* **15**, 299-302.
4. Kim, H. S., Y. Chen and P. Lonai. 1998. Complex regulation of multiple cytohesin like genes in murine tissues and cells. *FEBS Lett* . **433**, 312-316.
5. Kim, H. S., H. Hirai and O. Takenaka. 1996. Molecular features of the TSPY gene of gibbons and Old World monkeys. *Chrom. Res.* **4**, 500-506.
6. Kumar, S., K. Tamura and N. Nei. 1993. MEGA: Molecular evolutionary genetics analysis, version 1.01. The Pennsylvania State University, University Park, PA 16802.
7. La Mantia, G., D. Maglione, G. Pengue, A. Di Cristofano, A. Simeone, L. Lanfrancone and L. Lania. 1991. Identification and characterization of a novel human endogenous retroviral sequences preferentially expressed in undifferential embryonal carcinoma cells. *Nucleic Acids Res.* **19**, 1513-1520.
8. Leib-Mosch, C., M. Haltmeier, T. Werner, E. M. Geigl, R. Brack-Werner, U. Francke, V. Erfle and R. Hehlmann. 1993. Genomic distribution and transcription of solitary HERV-K LTRs. *Genomics* **18**, 261-269.
9. O'Connell, C., S. O'Brien, W. G. Nash and M. Cohen. 1984. ERV3, a full-length human endogenous provirus: Chromosomal localization and evolutionary relationships. *Virology* **138**, 225-235.
10. Shi, A., E. E. Coutavas and M. G. Rush. 1991. Evolutionary implications of primate endogenous retroviruses. *Virology* **182**, 495-502.
11. Tristem, M. 2000. Identification and characterization of novel human endogenous retrovirus families by phylogenetic screening of the human genome mapping project database. *J. Virol.* **74**, 3715-3730.

(Received June 29, 2001; Accepted August 10, 2001)

초록 : 인간 내생 레트로바이러스 HERV-S의 LTR엘리먼트의 동정과 계통분류

최주영 · 이주미 · 전승희 · 신경미 · 이지원 · 이원호 · 김희수*
(부산대학교 생명과학부)

최근 새로운 인간 내생 레트로바이러스 패밀리 (HERV-S)가 인간의 X 염색체상에서 동정 되었다. 그 길이는 6.7 kb 이며 LTR-*gag-pol-env*-LTR 의 일반적인 레트로바이러스의 구조를 가졌다. PCR 방법과 염기서열분석을 통하여 인간 게놈 DNA에서 HERV-S LTR패밀리를 동정하였다. 네 개의 LTR엘리먼트(HSL-1, HSL-5, HSL-10, HSL-11)가 동정 되었으며, 이들은 HERV-S LTR과 높은 염기배열의 유사성(96~99%)을 보였다. 계통분류학적인 방법으로 분석해보면 HERV-S LTR 패밀리는 영장류의 진화과정에서 진화적인 분기를 통해 주된 2개의 그룹으로 나뉘어졌다. 영장류에서 이러한 HERV-S LTR들의 연구가 이루어진다면 이들의 영장류 게놈 내의 삽입시기를 알 수 있고 또한 인류의 진화를 이해하는데 크게 이바지 할 것이다.